DOCUMENT RESUME

ED 140 346 CS 203 498

AUTHOR Day, Ruth S.
TITLE Verbal Fluency and the Language-Bound Effect.
INSTITUTION Yale Univ., New Haven, Conn. Dept. of Psychology.
SPONS AGENCY Office of Naval Research, Arlington, Va. Personnel and Training Research Programs Office.
REPORT NO CNR-TR-1
PUB DATE May 77
CONTRACT N0014-75-C-0967, NR-154-378
NOTE 46p.

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS *Cognitive Processes; College Students; Higher Education; *Language Research; Language Skills; *Memory; Phonetics; Psycholinguistics; Psychological Studies; *Verbal Ability; Visual Perception

ABSTRACT

Individuals previously identified as language-bound (LB) and language-optional (LO) participated in a series of experiments to study verbal fluency. The two groups showed a striking similarity in the number of responses they produced for categories with constraints at various levels (word form, word content, sentence, interpretation). This similarity occurred for both written and oral modes of response, and over a wide range of time intervals. Other types of measures, however, suggested that the forms in which a given category can be represented affected the ease with which the two groups produced their responses. LBs had more difficulty with categories that lent themselves readily to a spatial representation, while LOs had more difficulty with a category based on phonetic constraints. The results were considered in terms of their implications for the LB phenomenon, as well as general approaches to the study of verbal fluency. (Author)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED140346

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Technical Report #1
May, 1977

Verbal Fluency and the Language-Bound Effect

Ruth S. Day

This research was supported by:

Office of Naval Research
Psychological Sciences Division
Personnel and Training Research Programs
Contract Number: N0014-75-C-0967, NR 154-378
Scientific Officers: Dr. Marshall J. Farr and Dr. Henry M. Halff

Principal Investigator: Ruth S. Day
Department of Psychology
Yale University
New Haven, Connecticut 06520
(203) 432-4893

The views and conclusions contained in this document are those of the author and should not be interpreted as necessarily representing the official policies, either expressed or implied, by the Office of Naval Research or the U.S. Government.

Approved for public release; distribution unlimited

Reproduction in whole or in part is permitted for any purpose by the U.S. Government

CE 203 405

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: Technical Report No. 1	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): VERBAL FLUENCY AND THE ANGUAGE-BOUND EFFECT	5. TYPE OF REPORT & PERIOD COVERED: Interim Technical Report
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Ruth S. Day	8. CONTRACT OR GRANT NUMBER(s): N0014-75-C-0967
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Department of P ychology Yale University Box 11A Yale Station, New Haven, Ct. 06520	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR 154-378
11. CONTROLLING OFFICE NAME AND ADDRESS: Personnel and Training Research Programs Office of Naval Research (Code 458) Arlington, Virginia 22217	12. REPORT DATE: May, 1977 13. NUMBER OF PAGES: 40
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

This paper will appear in W. S.-Y. Wang and C. J. Fillmore (Eds.), Individual Differences in Language Ability and Language Behavior. New York: McGraw-Hill, 1978.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Individual differences; cognition; information processing; verbal fluency; levels of processing; phonetic; spatial; long-term memory; free recall; clustering; psycholinguistics; "language-bound;" "language-optional."

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

See reverse side.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Abstract

Individuals previously identified as language-bound (LB) and language-optional (LO) participated in a series of experiments designed to study verbal fluency. The two groups showed a striking similarity in the number of responses they produced for categories with constraints at various levels (word form, word content, sentence, interpretation). This similarity occurred for both written and oral modes of response, and over a wide range of time intervals. Other types of measures, however, suggested that the form(s) in which a given category can be represented affected the ease with which the two groups produced their responses. LBs had more difficulty with categories that lent themsleves readily to a spatial representation, while LOs had more difficulty with a category based on phonetic constraints. The results were considered in terms of their implications for the LB phenomenon as well as general approaches to the study of verbal fluency.

Abstract

Individuals previously identified as language-bound (LB) and language-optional (LO) participated in a series of experiments designed to study verbal fluency. The two groups showed a striking similarity in the number of responses they produced for categories with constraints at various levels (word form, word content, sentence, interpretation). This similarity occurred for both written and oral modes of response, and over a wide range of time intervals. Other types of measures, however, suggested that the form(s) in which a given category can be represented affected the ease with which the two groups produced their responses. LBs had more difficulty with categories that lent themselves readily to a spatial representation, while LOs had more difficulty with a category based on phonetic constraints. The results were considered in terms of their implications for the LB phenomenon as well as general approaches to the study of verbal fluency.

The proverbial person in the street, if asked to describe how individuals differ in "language ability," would probably say something about fluency. Some people talk easily while others have less to say and/or say it in a slower more halting fashion. This contrast in general verbal fluency is built into many of our expectations concerning different sorts of individuals. Say that we ask various people a simple question, such as "How many people work for you?" A talk-show host might be expected to say that more people have to work to get background information for certain types of quests and to negotiate terms for getting highly temperamental guests, while fewer people have to work when he knows the guest or conducts the negotiations himself, but then again, that all of this depends on how busy his own schedule has been in terms of other shows he is doing, travelling commitments, ...and on and on. In contrast, when a native from the state of Maine is asked how many people work for him, he might say something like "'bout half" (Starbird, 1977, p. 475).

What does verbal fluency reflect? Various studies have suggested that it reflects either intelligence or creativity (see Murphy, 1973, for a recent evaluation of these positions). However it could reflect in part the extent to which a person relies on linguistic as opposed to other forms of representation. For example, when people are asked to give as many states of the United States as possible in a brief amount of time, some might rely heavily on phonetic similarity to generate responses, as in MAINE, MONTANA, MICHIGAN, MISSISSIPPI, MISSOURI, while others might rely more on a "mental map" as in MAINE, NEW HAMPSHIRE, VERMONT, MASSACHUSETTS, RHODE ISLAND, CONNECTICUT. If a category specified in a fluency test can be represented in alternative forms, then two individuals could produce the same number of items in a limited period, yet do so in very different ways. However if the particular category used makes representation of the information easy in one form but not

others, and if people differ in the extent to which they rely on this form, then some individuals may be at a disadvantage.

Recent work suggests that individuals may indeed differ in the extent to which they rely on linguistic structures (Day, 1969; 1977). Some appear to be "language-bound" (LB): they perceive and remember events in language terms even when this approach leads them into misperceptions and distorted memories. Others appear to be "language-optional" (LO): they can use language structures or set them aside, depending on task demands. Classification of individuals as LB or LO is based on a temporal order judgment (TOJ) task involving fusible dichotic items. On a typical trial, an utterance such as BANKET is presented to one ear over earphones while LANKET is presented to the other ear. One of these items begins slightly before the other (by 50-125 msec) and the subject is asked to report "which sound" (phoneme) began first. LBs usually report hearing /b/ first even when /l/ led by a considerable interval; thus they report hearing only what their language allows, namely /bl/ in initial position but not /lb/. In contrast, LOs report the correct phoneme no matter which led; thus they can set aside linguistic rules concerning phoneme sequence and accurately perceive the events as presented. The LB effect is not based primarily on an effort to achieve meaningful percepts (as in BLANKET) since LBs still have difficulty with nonsense items such as BA/LA or GORIGIN/LORIGIN. However they can accurately judge temporal order when fusions can occur in either order (GAS/GAP can be fused into GASP or GAPS) or when no fusions are possible (BA/GA cannot be fused into either BGA or GBA). Thus LBs have trouble only when phonological rules of their language are violated by the temporal arrangements of the stimuli.

The LB-LO distinction extends beyond the domain of dichotic listening experiments. For example, the two groups perform differently in other auditory tasks, such as digit memory (Day, 1973a) and "secret language" transla-

tion (Day, 1973b), as well as visual word search (Day, 1974). The latter study is of particular interest for the present work. Subjects were asked to find words belonging to a particular semantic category which were embedded in a large matrix of letters. LBs and LOs did not differ in the number of words they found spelled out in the normal left-to-right direction (whether horizontal or diagonal). However LOs found more that were spelled out in right-to-left fashion and hence violated spelling conventions of English. In post-session interviews LBs typically reported that they "sounded out" letter sequences and then decided whether pronounceable strings fit the target category; either they did not, or could not, scan the matrix in all eight possible directions, even though they were told that words would be arrayed along all of them. LOs typically reported that they scanned the letters "visually" without a phonetic representation until acceptable items "lept out at them." Evidently, LBs relied more heavily on a phonetic representation of the matrix while LOs relied more heavily on a spatial representation. Thus, given a set of information that was presented spatially but could be represented in either spatial or linguistic forms, the two groups had clearly different preferences.

The present series of experiments was designed to determine whether LBs and LOs differ in a task that emphasized linguistic representation, namely verbal fluency. Certain test categories were included that could be represented in spatial as well as linguistic form. The ways in which subjects produced items was of interest, as well as the number of items that they could produce in a limited time period.

Method

Three experiments were conducted to study the verbal fluency of LBs and LOs. Although the experiments differed in some aspects, their general approach was very similar. Therefore it is useful to describe the methods used for all of them together.

Subjects

The 144 subjects were students from the introductory cognition course at Yale and met certain *a priori* criteria: they were right-handed, had no history of hearing trouble, and spoke English as their native language. Separate groups of students were studied in three experiments, drawn from different editions of the course taught over a four-year period. All were classified as LB or LO on the basis of the dichotic fusion TOJ task. In Experiment I there were 16 LBs (11 male, 5 female) and 21 LOs (16 male, 5 female); in Experiment II there were 26 LBs (14 male, 12 female) and 26 LOs (12 male, 14 female); and in Experiment III there were 25 LBs (16 male, 9 female) and 30 LOs (12 male, 18 female).

Category Constraints

The experiments differed in the type of constraints they placed on verbal production. Across all experiments, four general levels of constraints were studied, yielding ten categories in all, as shown in Figure 1. Experiment I required subjects to produce words to fit the following categories: Cities, Clothing, Flowers, Vegetables. These categories were originally intended to represent a semantic constraint. However since subjects often clustered their responses for Cities into foreign versus domestic cities, it was clear that this category involved a spatial constraint as well. In fact spatial considerations can conceivably be used to produce items for all of these

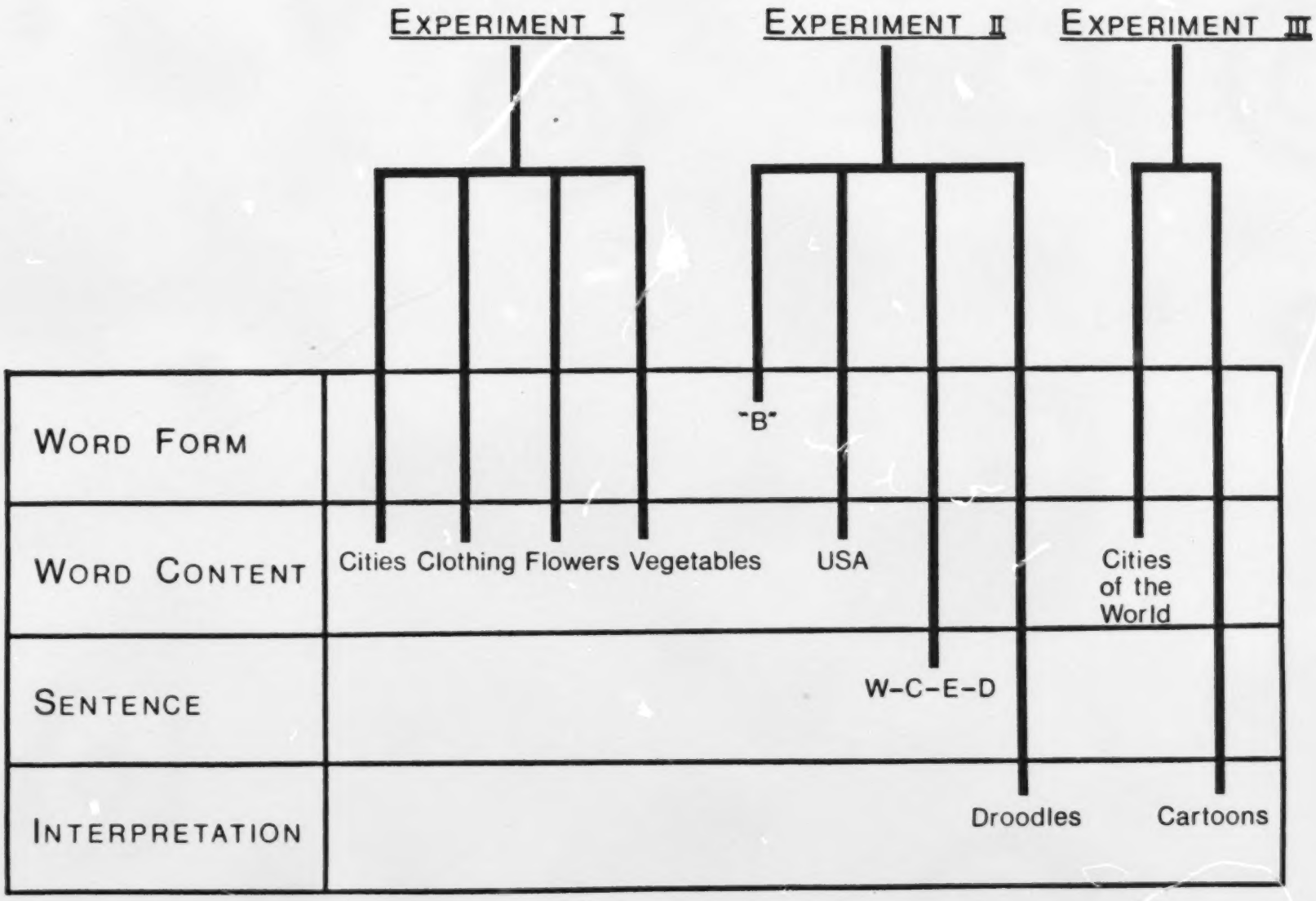

Figure 1 - Categories studied in each experiment, based on four levels of constraints.

"semantic" categories. For example, Clothing items can be partitioned according to the parts of the body on which they are worn, Flowers can be partitioned into those that grow singly from the ground versus those that grow on bushes, and Vegetables can be partitioned into those that go into the same "dish" such as a salad. Of course the spatial aspect is more salient in Cities than in these other categories. For the purposes of the present work, semantic and spatial considerations were kept together in order to represent a broad concept of word "content" as opposed to word "form." Experiment II probed all four levels of constraints: subjects produced items according to a word form constraint (words that begin with the letter "B"), a word content constraint (states of the U.S.A.), and a sentence constraint involving word form, word content, and syntactic factors (four-word sentences in which the first letter of each word was always W-C-E-D as in WASHINGTON CROSSED EVERY DELAWARE and WHEN CAN EDITH DANCE?). These categories will be referred to as "B," USA, and W-C-E-D. Subjects also gave interpretations for simple line drawings known as Droodles, selected from those used by Bower, Karlin, and Dueck (1975). The Droodles were drawn on 5" x 8" cards and are shown in Figure 2, along with a caption for each. The captions were not shown to the subjects. Experiment III contrasted a content constraint (Cities of the World) with another interpretation task, namely, devising a caption for a Cartoon. The Cartoon was taken from the New Yorker magazine (August 23, 1976, p. 25) and depicted a balcony terrace scene of a man in sunglasses stretched out on a lounge chair with his feet sticking up over the ledge and a woman peering cautiously at him from behind. The caption, which was "In some deep, ineffable way, Bob, you change when you put on your Earth shoes," was removed and only the drawing itself was used.

Procedure

In all experiments subjects were asked to produce as many items as possible that fit the target category, and to do so as rapidly as possible.

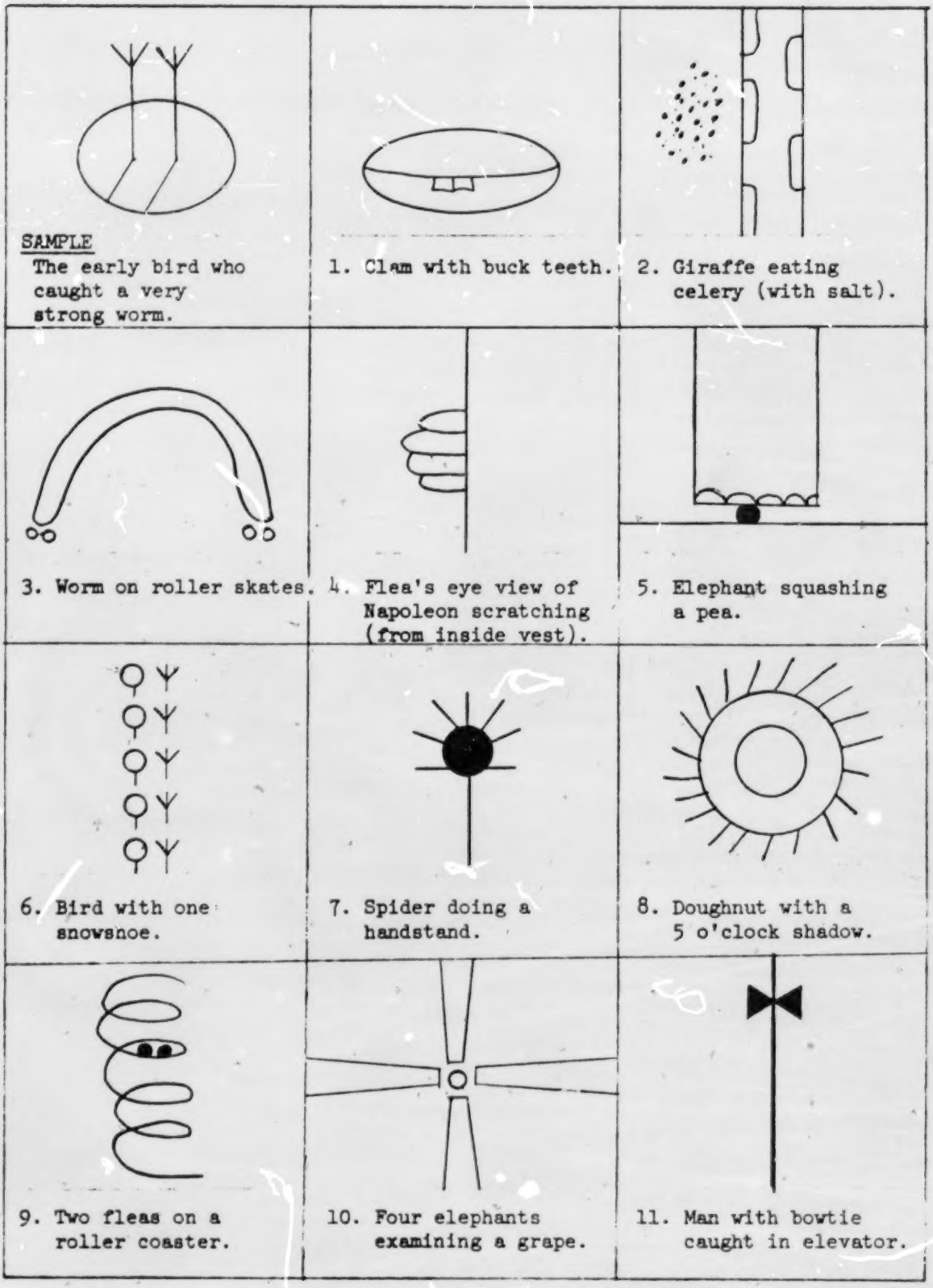
SAMPLE
The early bird who caught a very strong worm.
1. Clam with buck teeth.
2. Giraffe eating celery (with salt).
3. Worm on roller skates.
4. Flea's eye view of Napoleon scratching (from inside vest).
5. Elephant squashing a pea.
6. Bird with one snowshoe.
7. Spider doing a handstand.
8. Doughnut with a 5 o'clock shadow.
9. Two fleas on a roller coaster.
10. Four elephants examining a grape.
11. Man with bowtie caught in elevator.

Figure 2 - Droodles used in Experiment II. Items #1-11 served as stimuli and were shown without captions.

Mode of response. Subjects in Experiments I and III were tested in small groups and gave written responses on answer sheets with numbered blanks arranged in columns. Since subjects in Experiment II gave oral responses which were tape recorded, they were tested individually.

Time per category. Verbal fluency was only one of the concerns of an intensive testing program involving LB and LO subjects. Therefore the fluency tests were included along with other types of experiments to achieve sessions that were interesting yet not too taxing, so that the subjects would return for many additional sessions. The composition of these sessions varied across the three Experiments and hence different amounts of time were available for the fluency tests. In Experiment I five minutes were allotted for each category. Subjects were asked to circle the last item they had written at the end of the first 30 seconds, yielding two measured time intervals, 30 seconds and five minutes. In Experiment II, one minute was allotted for "B," USA, and W-C-E-D. The oral responses for "B" and USA were transcribed later onto answer sheets divided into 5-second intervals. Some analyses were performed for the successive 5-second intervals, however the data were also collapsed into 30-second and 1-minute units in order to make them comparable to the intervals used in the other experiments. Although the W-C-E-D data were also transcribed by 5-second intervals, subjects often produced long, irregular pauses between words or did not complete a sentence within five seconds; therefore data tabulations were prepared only for the full 1-minute interval. Only 15 seconds were allotted for each Droodle since pilot work showed that people tend to give an interpretation fairly quickly but then sit in silence thereafter. In Experiment III subjects wrote their responses to each category for three minutes and circled the last item written at the one- and two-minute points, yielding three measured intervals.

Counterbalancing. Counterbalancing conventions varied across experiments as a function of the mode of response and constraint levels studied. In Experiment I, since all categories were based on the same type of constraint and since it was one that generally elicits many responses, the order of categories was counterbalanced across subjects within each group (LB, LO). In Experiment II, an attempt was made to reduce the chance that subjects would become "tongue-tied" in giving oral responses in the presence of the tape recorder and experimenter. Thus the same order of categories was used for all subjects, beginning with those most likely to elicit many responses: USA, "B," W-C-E-D, and Droodles. Experiment III used the more "private" written response mode and pilot work suggested that it was easy to produce items for both Cities of the World and the Cartoon, so the order of the two categories was counterbalanced. However, unlike the previous experiments, Experiment III included the fluency tests in the initial screening session; since the status of each subject as LB or LO was not known at the time of testing, the counterbalancing could not be conducted systematically within each group.

Presentation of category constraints. In all experiments subjects were told to stop at the end of the allotted interval for each item. In Experiment I, they were also told when to turn the page for the next category, and the category name was printed at the top of each page. In Experiment II, the experimenter told the subjects aloud that the category was "states of the United States" and "words that begin with the letter 'B,' the second letter of the alphabet." A more detailed explanation was given for the sentence category along with some sample sentences for a different set of letters: "A-M-E-C" could be ALL MICE EAT CHEESE or ARISTOTELIANS MAKE EXCELLENT COMICS. Then they were told to use the letters, "W-C-E-D." Droodles also involved a somewhat more lengthy introduction. The sample Droodle shown in Figure 2 was shown first and two possible interpretations were given: THE EARLY BIRD WHO CAUGHT

A VERY STRONG WORM and BALLET DANCER IN A SPOTLIGHT. Subjects were asked whether they could "see" both interpretations and were told that Droodles have several possible interpretations with no one "correct" answer. Each card was then held in view for the allotted interval. In Experiment III subjects were told to write down "cities...located anywhere in the world," and that the cities should be "fairly well-known, recognizable to most well-educated people." For the Cartoon subjects were asked to write as many captions as possible and each had a copy of the Cartoon (without the caption) for the entire response interval.

Results and Discussion

General Approach

To date about 1,000 individuals have been studied in several forms of the dichotic fusion TOJ task and no sex differences in the language-bound effect have occurred. However, since women often score higher than men in tests of verbal fluency, sex was included as a factor in all analyses. Unless indicated otherwise, analyses of variance performed on the number of items produced used groups (LB, LO), sex, and cumulative time interval as factors, while those concerned with the proportion of clusters and other aspects of content used only groups and sex. Since the LB-LO distinction was of primary interest in these experiments, F-values are presented only for statistically reliable effects involving the groups factor and its interactions with other factors. All other reliable effects are reported without statistical information but were significant at the $p < .05$ level or better. Multiple comparisons were made using the Newman-Keuls procedure and all mentioned comparisons were reliable at $p < .05$ or better.

It was not possible to make formal comparisons across all categories since the three experiments used different configurations of time intervals, levels of category constraints, and task order conventions. Therefore separate analyses were performed for each of the ten categories. Experiment I did counterbalance categories within each group and therefore additional analyses were performed for it including category as a factor. Since there were some fairly gross differences across the ten categories, they are mentioned in the text in an informal way; however further experimentation would be needed to make these comparisons in a formal way.

Number of Items Produced

The primary measure in fluency tests is usually the number of items produced, which presumably reflects the ease with which subjects can "think of" items that fit a given category. There were no special problems in obtaining this measure in Experiments I and III since the written response mode yielded virtually no repetition of items by the same subject. However the oral response mode in Experiment II did yield 2% repetitions for "B" and 3% for USA. Subjects were usually aware that they had repeated an item as suggested by spontaneous comments such as, "Oops, I already said that," or their immediate repetition of an item to begin a new cluster, as in BOY, BABY, BALL...BALL, BASKETBALL, BEACHBALL. Analysis of the tapes showed that subjects were clearly aware of their repetitions most of the time (83% for each category). There were no differences between LBs and LOs in the number of repetitions they gave for "B." However LBs did give more repetitions (4%) than LOs (2%) for USA ($F(1,48) = 4.14$, $p < .05$). There were sizeable differences in repetition level between LB males (5%) and LO males (1%) but a comparable (3%) level for females in both groups ($F(1,48) = 4.18$, $p < .05$).

One way to view these repetition results is to assume that LBs and LOs ordinarily have comparable memory for items they have already produced, but that the presence of a potentially useful spatial component has different consequences for each group. LOs may rely more heavily on a mental map of the United States and somehow "cross off" each state as they say it, while LBs may form a less accurate mental image, have less careful spatial bookkeeping practices, or even dispense with mental images and rely more heavily on language-oriented means for producing states. The fact that females from the two groups did not differ in repetition levels for USA may be related to the fact that females generally score lower on measures of spatial ability such as Thurstone's Space Relations test; thus neither subset of females might be able to make special use of the spatial aspects of this task. In any event, repetitions were excluded from subsequent analyses.

An overview of the cumulative number of responses produced by LBs and LOs is shown in Figure 3. Casual inspection of the figure shows a striking similarity in the ease with which the two groups produced items, irrespective of the level of category constraint, the mode of response, or amount of time elapsed. Formal analyses showed that for nine of the ten categories there were no reliable differences between LBs and LOs. This finding did not vary as a function of either the sex of the subjects or the amount of time elapsed. For the remaining category, Cities of the World, LOs produced more items than LBs: 56.8 versus 50.0, respectively ($F(1,51) = 6.29$, $p < .05$). This difference between the groups increased as more time elapsed ($F(2,102) = 3.22$, $p < .05$). The fact that Cities (Experiment I) did not produce reliable group differences while Cities of the World (Experiment III) did is probably based on differences in instructions between the two tasks. In the first instance subjects saw just the single word, "Cities," as the definition of the category.

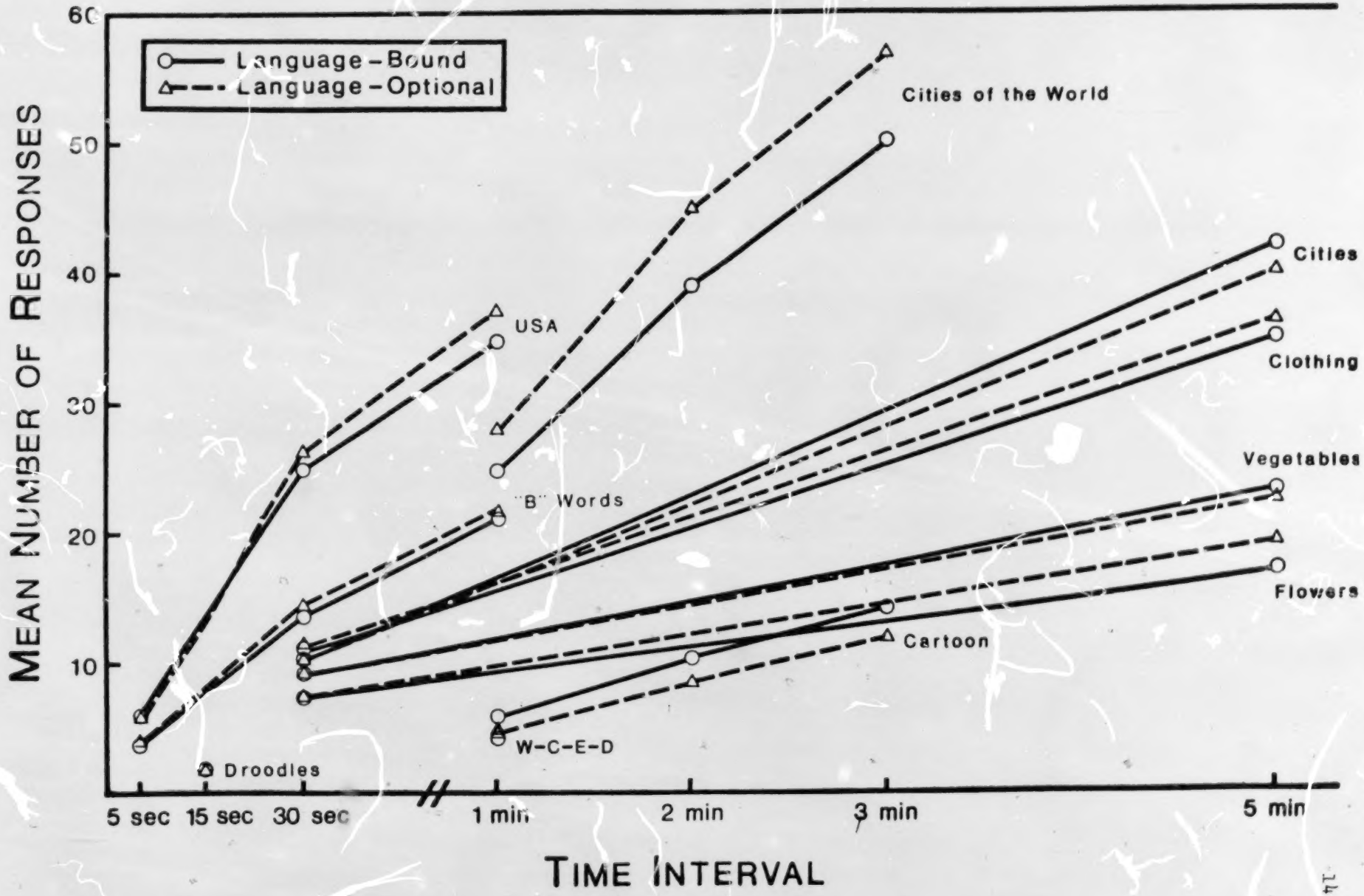

Figure 3 - Mean number of responses produced by language-bound and language-optional individuals for all categories and time intervals. Numbers are cumulative.

Since an analysis of the resulting protocols showed that some subjects gave primarily domestic cities while others gave many foreign cities as well, the category was extended to "Cities of the World" in order to determine whether a more explicitly articulated spatial domain would produce group differences. Also, since subjects in the Cities task sometimes gave low frequency items such as AVON, WEATOGUE, FARMINGTON, CANTON, SIMSBURY, which might indeed exist somewhere (in this case, in northwestern Connecticut) or be entirely fictitious, subjects in the later task were told to give only those cities which most well-educated people would be likely to know. Both of these new instructions widened the focus of the category and either or both may have been responsible for the greater number of responses produced by LOs. Nevertheless the overall findings of the three experiments are most notable for their lack of reliable differences between LBs and LOs.

There were reliable sex differences in four of the ten categories. Females produced more items for Clothing, Flowers, and Vegetables, while males produced more for Cities of the World. The shift to a male advantage in the latter case may well be based on the spatial aspect of the category. The differences between the sexes were more apparent at the longer time intervals for all of these categories except Vegetables.

There were substantial increases in the cumulative number of items produced in all categories as the amount of elapsed time increased. However, this effect was the typical one of decreasing gains as shown for the two categories that were analyzed by successive five-second intervals; by way of illustration, the mean number of items produced during the first and last five-second intervals was 4.1 and 1.3 for "B" and 5.8 and 0.9 for USA.

Category effects were studied formally in Experiment I since the order of the categories had been systematically counterbalanced within groups of LB and

LO subjects. The four word content categories did vary in difficulty with each reliably different from the others. This result suggests that the lack of LB-LO differences in number of items produced was not based on restricted "category width." Informal inspection of all the categories as displayed in Figure 3 suggests that it was easier to produce items based on word constraints than on sentence or interpretation constraints. Such results may not be based entirely on the speed with which people can "think of" items, since it takes longer simply to produce (write, say) phrases or sentences as opposed to single words. Nevertheless extending the time allotted might not yield substantial increases for categories involving sentence or interpretation constraints. This prediction is especially likely for Droodles since pilot subjects given unlimited time often produced an interpretation or two and then gave up. The extent to which W-C-E-D and the Cartoon would yield more responses with extended time depends partly on how one defines "different" responses; conceivably subjects could change a single word in a trivial way for each response and achieve an almost endless list of items, as in WHEN CAN EVA DRIVE, WHEN CAN EVIE DRIVE, WHEN CAN EDWARD DRIVE, WHEN CAN EDDIE DRIVE, WHEN CAN EDUARDO DRIVE.... or in WITH FEET LIKE YOURS IT'S A GOOD THING WE LIVE ON THE TENTH FLOOR, HARRY (GEORGE, HENRY, RUDOLPHO, WESTMORELAND....). Clearly, qualitative judgments must be considered as well as quantitative assessments in studying verbal fluency. The distinction between reproductive and productive processes is also important in studying fluency over extended time intervals. A category such as Flowers or "B" involves reproducing a limited set of items previously stored in memory, while the Cartoon involves producing a more open-ended (and perhaps unlimited) series of responses in a new setting. Therefore the slopes of the curves as shown in Figure 3 might level off more quickly for the word categories yet continue to rise for the Cartoon. Although there were

no statistically reliable differences between LBs and LOs in number of responses produced for the Cartoon, the slight advantage of LBs at each time interval coupled with the highly productive nature of this task suggests that it may be worthwhile to study responses to several Cartoons over longer time intervals. W-C-E-D responses might level off rapidly with extended time since it has so many built-in constraints.

Cluster Analyses

How does a person go about selecting individual words to fit a particular category such as Vegetables? He might rely on the sounds of words and try to find items that all start with the same phoneme, such as CUCUMBER, CORN, CAULIFLOWER, CABBAGE. Or, he might rely on subdivisions of content and try to find items that are all used in the same dish, such as the salad ingredients LETTUCE, TOMATO, CUCUMBER, CELERY. One way to determine whether such phonetic or content principles guided the subjects' search through memory is to examine the order in which they produced successive items.

Response protocols were examined for order effects in all seven word categories. Successive pairs of responses were scored as a phonetic cluster if they shared the same initial phoneme. Thus the sequence PARSNIPS, PEAS, POTATOES, SQUASH, BROCCOLI, BEETS has three phonetic clusters, two based on /p/ and one on /b/. Since responses in the "B" category already shared the same initial phoneme they were scored in terms of the following consonant (if any) and/or the first vowel. Thus BROOM, BLOOM, BLOSSOM contains two phonetic clusters, one based on /uw/ and one on /bl/.

It was more difficult to establish criteria for identifying content clusters. For example, Vegetables could be organized according to color; "leafy" versus "rooty;" usually eaten cooked versus raw; used primarily in Western, Eastern, Mid-Eastern, or other cookery; and so on. Some of these subcategories

demanded fairly sophisticated knowledge. Furthermore assignment of cases within even well-known subcategories was often troublesome. For example, what color is CABBAGE, red or green? In order to avoid as many of these problems as possible, only selected content analyses were performed. An exhaustive analysis was performed for "B" using all content subcategories that seemed reasonable. Thus all of the following items grouped together were considered to come from the same content subcategories: BABOON, BOBCAT, BADGER; BUBBLE, BURST; BARK, BITE; BISHOP, BERKELEY. The total number of clusters summed over all subcategories was obtained for each subject. A somewhat less exhaustive analysis was performed for Clothing using four subcategories: types of clothing (e.g., underwear), clothing with the same function (e.g., BELT, SUSPENDERS), articles worn on the same part of the body (e.g., head, torso, hands, feet), and articles worn for the same events (e.g., at bedtime, to formal affairs, for athletic activities, for cold weather). For Cities and Cities of the World, a single content criterion was examined, namely foreign (versus domestic) cities. USA responses were analyzed according to two spatial criteria. A stringent adjacency criterion required that the states share a common geographical border, while a lenient regional criterion required that they be from the same general region (Northeast, South, MidWest and Plains, or West).

The basic measure used in all cluster analyses was the proportion of clusters, which was evaluated as (Number of Clusters)/(Total Responses - 1) for each subject. Mean values for LBs and LOs are shown in Table 1 for the relevant categories. Again, the data are most notable for their overall lack of differences between LBs and LOs.

Phonetic clusters. In six of the seven analyses there were no reliable differences between LBs and LOs in the extent to which they relied on phonetic principles (as defined) to order their responses. In Cities of the World, which was described to subjects in a way that would facilitate spatial clustering,

Table 1 - Proportion of phonetic and content clusters in response protocols from categories based on word constraints.

	Cluster Type	
Categories	Phonetic	Content
Experiment I		
Cities	.07 (.08)	.25 (.22)
Clothing	.05 (07)	.46 (.51)
Flowers	.04 (.03)	---
Vegetables	.08 (.08)	---
Experiment II		
"B"	.20 (.17)	.19 (.25)*
USA	.09 (.14)	.44 (.47) Adjacency clusters .64 (.64) Regional clusters
Experiment III		
Cities of the World	.07 (.05)*	.28 (.30)

Note. The first number in each pair is the mean for LBs while the second (in parentheses) is for LOs.

*$p < .05$

LBs used phonetic principles more than LOs ($F(1,51) = 5.51$, $p < .05$). There were no differences between males and females nor any interactions between groups and sex in any of the analyses.

Content Clusters. Four of the five categories showed no reliable differences between LBs and LOs in the extent to which they relied on content principles to produce items. However LOs gave more content clusters for the "B" category ($F(1,48) = 4.52$, $p < .05$). Again, there were no differences between males and females in any of the analyses nor any interactions of group and sex.

Syntactic Repetition. The cluster analyses described above were designed to study the extent to which subjects relied on phonetic and content principles to produce single words. Sometimes subjects relied on syntactic principles in producing sentences for W-C-E-D. In the simplest case, they changed a single word in each successive sentence, as in WASHERS CLEAN EVERY DAY, WASHERS COOK EVERY DAY, thus retaining the same basic syntactic pattern. The same syntactic pattern can be retained in more adventurous ways, as by one subject who produced this sequence: WHITE CHEESE EXHIBITS DISTASTE, WRONG CORRECTIONS EASILY DISTRACT, WHITE CARNATIONS EAT DALMATIONS, WIDER COURSE-LOADS EXCLAIM DIRGES, WHOLE CARROTS EXUDE DETERMINATION, WORRIED CABINETMAKERS EXUDE DETERMINATION. Since subjects produced far fewer responses in W-C-E-D than for the word categories, a formal cluster analysis of syntactic patterns was not conducted. Instead a more lenient syntactic repetition score was obtained for each subject, based on the mean number of sentences he gave per syntactic pattern. There was no reliable difference between LBs and LOs in this measure; mean scores were 1.45 and 1.75, respectively.

Overview. Taken together, these clustering and repetition analyses suggest that LBs and LOs generally did not differ in the extent to which they relied on phonetic, content, or syntactic principles to generate items for the various categories. Interpretation of the two (out of 14) comparisons that

did yield reliable differences depends on how clustering and repetition scores are viewed. They could reflect a "better" representation for a given category. If so, then a group with a higher clustering score should also produce more items. Since this relationship did not occur, then a very different interpretation seems plausible. Greater reliance on clustering might instead reflect a "poorer" category representation in terms of number, strength, or availability of items, so that subjects must rely on such principles in a more explicit way in order to produce a reasonable number of items. Given this interpretation, then the results suggest that LBs had more difficulty with one of the categories involving a fairly strong spatial component (Cities of the World) and turned more heavily toward phonetic means for producing items for it, while LOs had more trouble generating words on a phonetic basis ("B") and turned more heavily toward content principles to compensate for this difficulty. The fact discussed previously that LBs also produced fewer responses for a spatial category (Cities of the World) and gave more repetitions for another (USA) supports this way of viewing clustering scores. (For a discussion of the possible roles that clustering might play in a related paradigm, see Crowder, 1976, pp. 329-333).

Commonality

Some responses were very common in a given category while others occurred infrequently. In order to determine whether LBs and LOs differed in the commonality of their responses, analyses were performed for selected word categories ("B," USA) and for all sentence and interpretation categories (W-C-E-D, Droodles, Cartoon). The first step was to construct a response hierarchy for each category containing all responses and their respective frequencies. Then a commonality score was computed for each subject, based on the mean frequency of all the items he produced.

It was a straight-forward matter to obtain frequency hierarchies for the word categories. The most common words given for "B" and their respective fre-

quencies were BALL(23), BOY(22), and BANANA(19), while unique responses included BANDERSNATCH and BEGONIA. Although LOs had a somewhat higher commonality score for this category (6.9 versus 6.1, respectively), the difference did not meet the conventional level of reliability ($F(1,48) = 2.97$, $p < .10$). The most frequent states in USA were CALIFORNIA(50), FLORIDA(49), and SOUTH CAROLINA(49), while the least frequent were NEBRASKA(25) and WYOMING(26). LBs had a higher commonality score for USA (39.6) than LOs (38.8) ($F(1,48) = 5.40$, $p < .05$). Commonality scores can be interpreted in the same general way as clustering scores. According to this view, a person would give many highly common items when he has a sparsely-populated category and/or a lack of facility in retrieving items from it in rapid fashion. If so, then the present findings suggest that although the two groups did not differ in the overall number of items they produced for these categories, LOs may have had more difficulty doing so given a phonetic constraint ("B") while LBs had more difficulty given a spatial constraint (USA). However given the partial marginality of the statistical evaluations, this interpretation must be made in a tentative way.

For W-C-E-D a frequency was obtained for every word and then the frequencies were summed for each sentence. The most common words for each letter were WHITE(21) and WE(21); CAN(69) and CAT(17); EAT(72) and EVERY(23); DAY(18) and DOUGHNUTS(13). Thus sentences varied widely in commonality scores, as in WHITE CATS EAT DOUGHNUTS (123) and WILD STRAWBERRIES EXCEED DOORSTEPS (7). There was no reliable difference between LBs (72.9) and LOs (64.1) in mean sentence commonality scores.

Representative Droodles (#1-6) were scored in terms of the commonality of the main object in each drawing. The response hierarchies are given in Table 2 in order to show the wide range of object interpretations that occurred despite the very simple nature of the line configurations. Some liberties were taken in obtaining the frequencies shown in the table. For example, in Droodle #1,

Table 2 - Commonality of the main object in each Droodle. Numbers indicate the frequency with which each object was given over all responses obtained. Note: objects named only once ("unique") are not listed separately although they are included in the totals.

Object	Frequency
Droodle #1	
1. Football	42
2. Mouth	14
3. Clam	13
4. Hamburger	12
5. Vehicle	12
6. Egg	9
7. Wash on line	7
8. Creature	3
9. Dish	2
10. Unique responses (N=10)	1
	124
Droodle #2	
1. Giraffe	30
2. Street	14
3. Tree	12
4. Snake	9
5. Building	5
6. Microscope scene	5
7. Pole, post	4
8. Unique responses (N=9)	1
	88
Droodle #3	
1. Snake, worm	19
2. Rainbow	18
3. Arch	8
4. Ribbon	7
5. Creature	6
6. Headphones	4
7. Hotdog	2
8. Handle	2
9. Unique responses (N=13)	1
	79
Droodle #4	
1. Hand	51
2. Bird	5
3. Indian	3
4. Hotdogs	3
5. Cactus	3
6. Bananas	2
7. Tombstones	2
8. Unique responses (N=19)	1
	88
Droodle #5	
1. Elephant	49
2. Window shade	11
3. Curtain (on stage)	10
4. Something on ball	7
5. View through window	6
6. Seats in movie theater	3
7. Trapeze	2
8. Unique responses (N=7)	1
	95
Droodle #6	
1. Trees	38
2. Birds	20
3. Lollipops	15
4. (Unidentified) tracks	3
5. Dancers	2
6. Balloons	2
7. Flowers	2
8. Peace signs	2
9. Unique responses (N=3)	1
	87

GRINDER was tabulated with HAMBURGER, while SPACESHIP, FLYING SAUCER, BLIMP, ZEPPELIN, and SUBMARINE were all tabulated as VEHICLES. In some cases a single object interpretation predominated (e.g., HAND for Droodle #4), while for others several interpretations were fairly common (e.g., TREES, BIRDS, and LOLLIPOPS for Droodle #6). Separate analyses were performed on the object commonality scores for each Droodle. There were no differences in these scores for LBs and LOs in any of the six Droodles. Mean scores for each group, respectively, were 21.3 and 20.7 for Droodle #1; 16.6 and 17.2 for Droodle #2; 8.9 and 10.9 for Droodle #3; 33.9 and 35.3 for Droodle #4; 33.2 and 30.2 for Droodle #5; and 26.2 and 21.2 for Droodle #6. There was an interaction of group with sex for Droodle #3, with the following mean scores for each subgroup: 10.8 for LB males, 6.9 for LB females, 9.1 for LO males, and 12.7 for LO females ($F(1,48) = 4.36$, $p < .05$). Inspection of the response protocols gave no clues concerning the basis of this finding.

Captions for the Cartoon were scored in terms of the commonality of their general topic. The most frequent topic (reflected in 115 of the 727 captions) was "interruptions" in which the woman told the man that he has a phone call (usually from his mother or boss), that guests or relatives just arrived, that dinner is ready (or burned), and so on. The remaining topics and their frequencies for the most common topics were: "lazy bum"(45), "good life"(40), "drink"(38), "man's appearance"(34), "please fix the...."(26), "sunburn"(26), "what are you doing?"(25), "do you want anything"(21), "view, weather"(21), "pollution, traffic"(18), and "hank-panky"(16). Since the sorting of captions into general topics involved many arbitrary decisions, the basic measure used in the statistical analysis was the proportion of sentences that subjects produced based on these most common topics, rather than a mean commonality score. LBs and LOs did not differ in this measure (.67 and .68, respectively).

The results of all the commonality analyses taken together are somewhat mixed. Clearly, neither group relied more heavily on common responses in

order to interpret pictures or produce sentences. However, production of words may have involved systematic differences in commonality between the groups. As suggested tentatively above, if higher commonality scores are viewed as reflecting a compensation for a poorer category representation, then LBs had more trouble with a clearly spatial category (USA) while LOs had somewhat more trouble with a phonetic category ("B"). This interpretation complements that given for the cluster analyses in which LBs had (in some sense) more trouble with another heavily spatial category (Cities of the World) while LOs had more trouble with the same phonetic category ("B").

Quality of Responses

Despite the fact that LBs and LOs showed no gross differences in verbal fluency, they could still differ in the "quality" of the responses they gave. While quality is inevitably a difficult matter to assess, nevertheless some responses clearly seemed better than others in the sentence and interpretation tasks. Two general approaches were taken to assess the quality of responses in W-C-E-D, Droodles, and the Cartoon. Subjective assessments were made by raters who judged the "goodness" of the responses; the actual terms used to reflect goodness were varied in order to take into account the special problems posed by each task. More objective assessments were made of specific aspects of each task that appeared to contribute to the subjective measures of goodness.

W-C-E-D. The 15 new students who served as raters were told about the origin of the sentences and were asked to produce some of their own sentences for the same constraints. Then they rated the "overall creativity or inventiveness" of all sentences produced by the LB and LO subjects, using a 5-point scale ranging from 5 = "very creative" to 1 = "not creative." The sentences that each subject produced were listed in order on the rating sheets, but the order of subjects was random. Sample sentences and their mean creativity ratings were: WILL CORRUPTION END DYNASTIES? (3.7), WILLIAM CHASES EVERYONE'S DAUGHTER (2.8), WATER CAN ENTER DRIVEWAYS (2.3), WHICH CAT EYJELLS DUMBLY? (1.8)

and WHY COME EAT DOGS? (1.5). There was no reliable difference in mean creativity ratings for LBs (2.1) and LOs (2.2).

Raters were asked to establish their own criteria for rating the creativity of W-C-E-D responses and to use these criteria in a consistent way. In postsession descriptions of their approaches, many of them said that their evaluation was based in part on whether the responses were "good English sentences," although they did not define this term very clearly, if at all. Subsequent analyses were conducted to identify sentences that violated normal English in terms of syntactic, semantic, and pragmatic conventions. For example, WILL CANDY ENTER DOOR clearly violates English syntax. The sentence WHY CAN'T EVERYTHING DRINK? violates a semantic convention in that EVERYTHING and DRINK assume different types of entities: EVERYTHING implies living and nonliving objects, including plants, while DRINK implies something that living organisms do, usually animals. Of course in more poetic usages most anything could DRINK, as in THE WET BARN DRANK IN THE RAYS OF THE SUN. Nevertheless if ordinary semantic agreement rules were violated, the sentence was scored as semantically unacceptable. The sentence, WOMEN CARRY EXCELLENT DOGS raises all sorts of pragmatic questions, for example, Why do they carry only excellent dogs? What is the difference between excellent and nonexcellent dogs? Why do they carry dogs at all, anyway? Of course one can imagine a fantasy world in which dogs are the rulers and differ in social status ("excellence") according to breed, while among humans women have higher status than men, such that women have the great privilege of carrying the more exalted dogs in stately processions. No doubt any sentence can be made to "make sense" given a fantasy world or metaphorical interpretation. However if unusual circumstances must be fashioned in order to comprehend it, then it violates ordinary pragmatic conventions. Each complete sentence in W-C-E-D was scored as acceptable or unacceptable in terms of syntactic, semantic, and pragmatic conventions. The proportions of unacceptable sentences varied systematically over these three types of violations, respectively: .04, .18, .37 ($F(2,96) = 47.11$, $p < .001$). LBs and LOs

differed only in the proportion of sentences which were unacceptable on pragmatic grounds ($F(2,96) = 3.38$, $p < .05$). Means for these violations, respectively, were .04, .17, .45 for LBs and .03, .18, .30 for LOs. There are contrasting ways to view the LBs' greater use of pragmatic violations. They may have had trouble thinking up sentences and therefore violated pragmatic rules out of desperation. Or, they may have used the task as an opportunity to be inventive (at least to themselves if not to the raters), and to deliberately violate normal language conventions. It would be interesting to know whether LBs can more readily find interpretations for expressions involving unusual language usage. To some people, the string, WILD STRAWBERRIES EXCEED DOORSTEPS simply does not make sense. Others, however, understand that certain strawberries are so wild that they even grow up over man-made objects such as doorsteps. Some people may understand more readily that WOMEN CAN EAT DIFFERENCES, that WHITE COLORS EXHIBIT DIETING, and know when to ask, WAS CHRISTMAS EVERY DAY? Is the ability to obtain interpretations for such sentences related to the LB phenomenon? The raters clearly differed in their evaluations of such sentences; some gave them consistently low ratings while others gave them high ratings. Unfortunately the status of raters as LB or LO was not known.

Droodles. The five new students who rated Droodles captions began by studying the sample Droodle and its caption (as shown in Figure 2). Then they examined each Droodle without its caption and rated the "goodness" of the captions produced by LBs and LOs along a 5-point scale from 5 = "excellent" to 1 = "poor". They were told to consider the following criteria in judging goodness: whether the caption accounted for all or only part of the figure and did so in an integrated rather than enumerative way; whether it was reasonable, interesting, or amusing; and whether it was in some way abstract rather than a literal description of lines. The same six drawings that were examined in the commonality analyses were rated, namely, Droodles #1-6. This was a fairly demanding task as the raters had to evaluate a total of 563 captions. Examples of cap-

tions and their mean ratings for Droodle #1 are A HAMBURGER THAT SWALLOWED A GOPHER (5.0), TWO SHEETS HANGING ON A LINE, SEEN THROUGH A PORTHOLE (4.8), FOOTBALL (1.2), and AN EASTER EGG (1.2). Examples for Droodle #3 are MOBILE McDONALD'S ARCH (4.6), RAINBOW WITH FOUR POTS OF GOLD (4.2), and TOP OF SOMEONE'S HEAD (1.4). LOs produced reliably better captions for three of the six Droodles. The mean goodness ratings for LBs and LOs, respectively, were 2.7 and 3.5 for Droodle #1, ($F(1,48) = 19.84$, $p < .001$); 2.3 and 2.9 for Droodle #3 ($F(1,48) = 4.87$, $p < .05$); and 3.0 and 3.5 for Droodle #5 ($F(1,48) = 7.25$, $p < .01$). In two of these Droodles there was a similar interaction of groups and sex. Basically, it suggested that the LB-LO distinction is more important for males than for females, at least in producing "good" Droodle captions; for Droodle #1, the mean goodness scores for LB males, LO males, LB females and LO females, respectively, were 2.5, 3.7, 2.9, and 3.3 ($F(1,48) = 6.14$, $p < .05$), while for Droodle #5 these values were 2.4, 3.8, 3.5, and 3.2 ($F(1,48) = 17.21$, $p < .001$). In Droodles #2, 4, 6, LOs had numerically higher goodness scores than LBs but these differences were not reliable.

Most of the Droodles yielded descriptions viewed from a single perspective. For example, captions for Droodle #1 indicated that the viewer was taking a side view, i.e., that he was positioned perpendicular to the exposed part of the figure. However Droodle #2 had at least two perspectives: from the side, as in TREE WITH A SWARM OF INSECTS BY IT, or from above, as in STREET WITH CARS PARKED ALONG IT AND BUSHES ON ONE SIDE. There were about twice as many captions from the side view (52) as from the aerial view (28). The perspectives of seven remaining captions for this Droodle were unclear as in BEES--THE LITTLE DOTS, while in one it was from a vantage point rotated 90°: SIDEWAYS VIEW OF A PERSON'S MOUTH WITH A BEARD UNDER IT. Viewing a drawing from a particular perspective may make it more difficult to view it from another perspective. Protocols were examined to determine whether the captions

produced by each subject for Droodle #2 were based on only one or on both perspectives. LBs had a higher mean number of perspectives (1.4 versus 1.2, respectively, $F(1,48) = 4.46$, $p < .05$). The finding is somewhat surprising since the groups did not differ in goodness ratings for this Droodle and furthermore LOs had significantly higher goodness scores on half the Droodles. Since the two perspectives differed in their relative frequencies of occurrence, LBs still might have relied more heavily on the more common side view. All responses were given a perspective commonality score and then means were obtained for individual subjects. There was no reliable difference in mean perspective commonality scores between LBs (40.2) and LOs (41.3), although females did score higher than males (44.5 versus 37.0, respectively). Once again, if higher commonality scores are viewed as reflecting less facility in a task, then the latter result suggests that females have less facility in spontaneously viewing the Droodles from unusual perspectives.

In most of the Droodles there was little variation in figure-ground relationships. However the interpretation of the vertical line in Droodle #4 led to very different configurations which sometime included changes in perspective. The line could establish: a large figure such as a wall on the right with smaller objects protruding from behind it, as in GUY CARRYING PILLOWS and PARROT ESCAPING DOWN THE STREET; a large figure on the left with smaller objects in front of it, as in A JELLO MOLD THAT SOMEONE THREW AGAINST THE WALL AND IT STUCK; a continuous figure with some sort of seam in the middle, as in PERSON PUTTING HANDS FROM OCEAN LAST TIME DOWN AS DROWNING; a very thin figure in the middle as in FLAGS ON LINE; or, if rotated 90°, a small figure perpendicular to a large one as in A CACTUS ON THE HORIZON. Sometimes, however, it was not possible to determine the figure-ground relationship; for example, the caption FINGERS REACHING AROUND A CORNER could refer to a hand with fingers

already bent and knuckles showing in front of a wall on the left, or a hand not yet bent with the insides of the fingers shown coming from behind a wall on the right. Ability to spontaneously see several figure-ground configurations suggests a type of spatial flexibility. If so, then LBs and LOs showed comparable spatial flexibility in Droodle #4, since both groups produced the same mean number of configurations (1.5). Analysis of these data in terms of perspective commonality (as described for Droodle #2) yielded no reliable differences between the groups either (22.6 versus 21.9 for LBs and LOs, respectively). Thus the general advantage of LOs in producing "good captions" was not based on the ability to take alternative perspectives or form alternative figure-ground relationships.

Cartoon captions. Most of the Cartoon captions were quite pedestrian, as in GORGEOUS DAY, DINNER'S READY, and WANT ANOTHER DRINK DEAR? Some however, were possibly funny, as in IF I'D MARRIED A TALL MAN WE'D HAVE TO GET A LARGER TERRACE; WHEN HE ASKED ME TO SLIP INTO SOMETHING MORE COMFORTABLE I DIDN'T KNOW HE MEANT A DEEP SLEEP; FOR THIS YOU GET $30,000 A YEAR AS A CITY PLANNER?; WHO DREW US?; YOU'LL NEVER BELIEVE IT BUT OUR SON WANTS TO GO TO YALE; IF JERRY BROWN WINS WILL WE HAVE TO GIVE ALL OF THIS UP? Instead of having many judges rate the captions by reading each once, only one highly experienced viewer of New Yorker cartoons judged them by studying all 727 captions several times and then selecting those that seemed "funny" enough to appear in the New Yorker. LBs and LOs did not differ reliably in the proportion of humorous captions they produced (.12 and .18, respectively).

Sometimes the use of very specific terms increased the cleverness of the captions, as in CAPTAIN KANGAROO JUST CAME ON--DO YOU WANT IT IN COLOR as opposed to COME IN DEAR, THE TV SHOW IS STARTING; or DOES THAT CLOUD LOOK TO YOU LIKE CARAVAGGIO'S BACCHUS as opposed to IT LOOKS LIKE RAIN. Each sentence

was scored for presence of specific terms (excluding the man's name which was usually George or Harry, and items to be fixed which were often lawnmowers or toilets). LOs gave a higher proportion of sentences with specific terms than LBs (.30 versus .14, respectively, $F(1,51) = 11.51$, $p < .01$). Since LBs did not produce a lower proportion of humorous captions than LOs, evidently they used some other means of being funny. Perhaps they relied more on a general conception of the scene rather than a more surface use of specific items. Males also had a higher proportion of humorous captions than females (.28 versus .16, respectively, $F(1,51) = 5.37$, $p < .05$).

Most captions were comments made from the woman's perspective (STOP OM-ING HAROLD, THE MCFINTONS ARE HERE). A few were from the man's perspective (WILL YOU GO BACK INTO THE HOUSE AND STOP NAGGING ME), while some could have come from either character (AH, HIGH CLASS LIVING IN QUEENS). Perhaps the most interesting set of captions came from the perspective of a third party, as in EXISTENTIAL ENNUI IN THE MODERN CITY and CAPITALIST WARMONGERS IDLE ABOVE THE PROLETARIAT. LBs and LOs gave a comparable proportion of their responses from the third person perspective (.22 and .14, respectively).

Overview. Assessments of response quality yielded clear differences between LBs and LOs and perhaps the most interesting findings in the study. LOs clearly gave better captions for the Droodles. This task requires interpretation of spatial relations and hence the superiority of LOs here complements the previous findings based on number of items reported, clustering, and commonality. Although LBs and LOs did not differ in the rated quality of their responses for W-C-E-D or the Cartoon, other analyses suggested that the two groups may have achieved these response levels in different ways. In W-C-E-D, LBs violated more pragmatic conventions; the resulting sentences require more conceptualization (e.g., of surrounding circumstances or

metaphorical interpretations) in order to comprehend them. In producing Cartoon captions, LBs gave fewer specific terms in order to achieve cleverness which suggests that they may have relied more heavily on a "deeper" conceptualization of the scene rather than on a more "surface" selection of terms. While there is little question concerning the spatial superiority of LOs in producing good Droodle captions, the suggestion concerning a possible greater "depth of conceptualization" for LBs in the other tasks must be regarded with considerable caution until further research is conducted to study this issue in a more explicit fashion.

General Discussion and Conclusions

Basic Findings

It is sometimes a delicate matter to know when to stop doing research on a particular question, both in terms of the number of experiments and types of analyses to conduct. In the present case, work might well have been ended after Experiment I for it failed to show fluency differences between LBs and LOs in any of the categories studied. Or, it might have been ended after examining the number of items produced across all three experiments since nine of the ten categories studied failed to yield LB-LO differences, despite the fact that various levels of category constraints and modes of response were used. However the nature of the one category that did yield group differences in number of responses produced suggested that the form(s) in which a set of information can be represented may be an important consideration in studying individual differences in verbal fluency. LBs produced fewer items for Cities of the World which can be represented in spatial as well as linguistic terms. They also had more trouble remembering which items they had already produced in an oral task involving another heavily spatial category, USA. Both of these results suggested that LBs may be so heavily dependent on linguistic representation that they have difficulty using other forms of representation even when they are useful.

Analyses of content suggested that LBs and LOs sometime achieved their (usually) comparable response levels in different ways. While many of these analyses yielded no reliable differences between the groups, those that did occur were generally interpretable in terms of a linguistic-spatial distinction. There are various ways that subjects can try to increase the number of items they produce. For example, they can rely on clustering principles to generate successive items or they can give highly common items. Both groups clearly

relied on such principles at least to a certain extent. However in those cases where one group relied on such an approach more heavily yet did not achieve a greater number of responses, then their reliance can be viewed as a way to compensate for a poorer category representation. If so, then LBs had more difficulty in representing categories with clear spatial aspects since they gave a higher proportion of (phonetic) clusters for Cities of the World and more common items for USA. By the same argument, LOs had more difficulty in representing a category based on a phonetic constraint since they gave a higher proportion of (content) clusters for "B" and also had a tendency to give more common words for this same category. Furthermore, subjective measures of response quality in the sentence and interpretation tasks showed group differences only when there was an especially heavy spatial aspect to the task; thus LBs produced poorer captions for the Droodles yet did not differ from LOs in producing good sentences for W-C-E-D or captions for the Cartoon.

The phonetic-spatial results complement those obtained in a previous word-search task (Day, 1974) in which LBs relied more heavily on phonetic representation and LOs on spatial representation. The two sets of studies provide an interesting contrast since the word search experiment presented information in spatial form while the fluency experiments emphasized linguistic representation; nevertheless LBs and LOs retained their preferences concerning representational form in both sets of circumstances.

Many investigators have shown that phonetic processing takes place primarily in the left cerebral hemisphere of the brain (e.g., Studdert-Kennedy and Shankweiler,1970) while spatial processing takes place primarily in the right hemisphere (e.g., Kimura, 1967). The present results suggest that LBs and LOs may differ in the extent to which they rely on these two types of brain systems. A subsequent paper presents data concerning hemispheric asymmetry in LBs and LOs.

Other aspects of the fluency data suggested that LBs may rely on "deeper" forms of information processing in some of the tasks, while LOs use more "surface" forms. However at present these suggestions are at best tentative.

Quantitative versus Qualitative Differences

Intensive study over the past few years has suggested that the LB-LO distinction is not based on quantitative differences in general intelligence level. For example, the two groups did not differ in a standard test of intelligence (Thurstone's Primary Mental Abilities Test) nor in the verbal or quantitative subtests of the Scholastic Aptitude Test. Instead, the differences appear to be based on qualitative differences in reliance on language structures. Although there has been considerable controversy in the literature concerning whether "verbal productive thinking" reflects intelligence or is independent of it (see Murphy, 1973), the present work is still informative concerning the quantitative-qualitative distinction. Most of the present data did not yield reliable differences between the groups in terms of the quantity of responses they produced. Instead, the differences that did occur were based almost wholly on more qualitative aspects of the data, namely, clustering, commonality, and response goodness. Thus the two groups often achieved the same quantitative level in qualitatively different ways.

Possible Implications for the Study of Verbal Fluency

Had the present experiments used only categories at the word content level that are amenable primarily to linguistic representation, then the performance of LBs and LOs would appear to be virtually indistinguishable. Tests typically used to study fluency (see Guilford and Hoefpner, 1971, for a recent approach) sometimes vary in their need for linguistic representation. Some do benefit from other forms of representation (as in giving alternative uses for a brick)

but many rely heavily on explicitly linguistic representations. Therefore if a category used in a fluency test is selected without regard for the nature of the forms that best represent it, then the results could reflect the reliance that individuals place on the particular form involved, rather than more general fluency abilities. Furthermore, since there appear to be more LBs than LOs in the general population, then the norms for standardized tests involving fluency may favor individuals who habitually rely heavily on linguistic representations and penalize those of comparable intellectual ability who work better with spatial representations.

Comprehensiveness of the Present Study

It could be argued that fluency has been studied here in a fairly narrow way. We can probably all recognize people who are fluent from those who are not in everyday life, but we do not ordinarily do so by asking them to produce as many words as possible in a brief amount of time. While it might be more "ecologically valid" to have subjects just talk freely, such an approach has many inherent methodological difficulties. Given the great amount of work needed to analyze the present highly-controlled data and the fairly consistent findings obtained, an exhaustive study of fluency might prove to be more exhausting than enlightening.

The present findings have not been discussed in terms of the vast amount of fluency research reported in the psychometric literature (see Horn, 1976, for a recent review). In the temporary absence of data from LBs and LOs based on the many types of tests used in psychometric studies, it seems best to defer that discussion to a later time. Among other things, however, it will be interesting to determine whether LBs and LOs differ in various aspects of fluency, such as expressional, associational, and ideational fluency. Murphy (1973) recently suggested that it may be useful to distinguish between "figural" and

"verbal" subfactors of verbal productive thinking; the present results underscore the usefulness of that idea.

Cautions Concerning Terminology

The terms "language-bound" and "language-optional" should be used with some caution. They might seem to imply that people so designated should perform differently in all tasks involving language processes. Such a view can be misleading. "Language ability" is composed of many aspects, and while these aspects may be intercorrelated, there is ample room for individuals to possess different configurations within this very broad ability. Therefore the two groups might well differ in some language tasks but not others. Since language is incredibly "overlearned," LBs and LOs need not show differences on all language tasks, especially if they involve processes that are highly overlearned. In any event, caution should be exercised in order to avoid becoming language-bound concerning the LB and LO terms themselves.

Conclusion

In conclusion, LBs and LOs do not appear to possess quantitative differences in verbal fluency. However the nature of the constraints imposed in the fluency tasks may cause subtle differences in response quality, such that LBs have more difficulty with categories based on spatial constraints while LOs have more difficulty with those based on phonetic constraints.

Acknowledgements

A number of people contributed to this work in both formal and informal ways. Among those who devoted considerable effort were Jane Weissman who transcribed the tapes in Experiment II and provided many preliminary data tabulations, Lynne Alvarez who conducted most of the statistical analyses, Cynthia Joyce who helped in many different ways, and Ray Jordan who was always standing by.

References

Bower, G.H., Karlin, M.B., and Dueck, A. Comprehension and memory for pictures. Memory and Cognition, 1975, 3, 216-220.

Crowder, R.G. Principles of Learning and Memory. Hillsdale, N.J.: Erlbaum, 1976.

Day, R.S. Temporal order judgments in speech: Are individuals language-bound or stimulus-bound? Paper presented at the 9th meeting of the Psychonomic Society, St. Louis, November, 1969. Also in Haskins Laboratories Status Report, 1970, SR-21/22, 71-87.

Day, R.S. Digit-span memory in language-bound and stimulus-bound subjects. Journal of the Acoustical Society of America, 1973(a), 54, 287(A). Also in Haskins Laboratories Status Report, 1973, SR-34, 127-139.

Day, R.S. On learning "secret languages." Paper presented at the Eastern Psychological Association meeting, Washington, D.C., May, 1973(b). Also in Haskins Laboratories Status Report, 1973, SR-34, 141-150.

Day, R.S. Differences between language-bound and stimulus-bound subjects in solving word search puzzles. Journal of the Acoustical Society of America, 1974, 55, 412(A).

Day, R.S. Systematic individual differences in information processing. In Research Frontiers section of Zimbardo, P.G. and Ruch, F.L. Psychology and Life. Glenview, Ill.: Scott, Foresman, 1977. Pp. 5A-5D.

Guilford, J.P. and Hoepfner, R. The Analysis of Intelligence. New York: McGraw-Hill, 1971.

Horn, J.L. Human abilities: A review of research and theory in the early 1970s. Annual Review of Psychology, 1976, 36, 437-485.

Kimura, D. Dual functional asymmetry of brain in visual perception. Neuropsychologia, 1967, 4, 275-285.

Murphy, R.T. Investigations of a Creativity Dimension. Princeton, N.J.: Educational Testing Service, RB-73-12, 1973.

Studdert-Kennedy, M. and Shankweiler, D.P. Hemispheric specialization for speech. Journal of the Acoustical Society of America, 1970, 48, 579-594.

Starbird, E. A way of life called Maine. National Geographic, 1977, 151, 727-757.